

sample is extracted from the ampoule and cleaned of the surrounding excess impregnant by standard mechanical polishing techniques. In this way, we prepared nanowires of various metals (In, Sn, and Al) and semiconductors (Se, Te, GaSb, and Bi<sub>2</sub>Te<sub>3</sub>) (Fig. 2).

The nanowire composites create substantial electric field patterns over the sample surface. We used a scanning probe microscope to measure electric fields at the surface of a nanocomposite. In a NanoScope (Digital Instruments, Santa Barbara, California) scanning force microscope, the sample is mounted with conductive epoxy to a metal holder and is held at a few volts relative to a conductive cantilever tip that is grounded. The metal-coated, etched, single-crystal silicon tip has a radius of curvature of about 5 nm. The tip is set to oscillate at a frequency near its resonance frequency (78 kHz). When the cantilever encounters a vertical electric field gradient, the effective spring constant is modified, shifting its resonance frequency. By recording the amplitude of the cantilever oscillations while scanning the sample surface, we obtain an image that reveals the strength of the electric force gradient (13, 14).

The image, however, may also contain topographical information; it is difficult to separate the two effects. This is circumvented by taking measurements in two passes over each scan line (15). On the first pass, a topographical image (Fig. 3A) is taken with the cantilever tapping the surface, and the information is stored in memory. On the second pass, the tip is lifted to a selected separation between the tip and local surface topography (typically 20 to 200 nm), such that the tip does not touch the surface. By using the stored topographical data instead of the standard feedback, we can keep the separation constant. In this second pass, cantilever oscillation amplitudes are sensitive to electric force gradients without being influenced by topographic features (Fig. 3B). This two-pass measurement process is recorded for every scan line, producing separate topographic and electric force images. From these images, contours of electric force gradient (Fig. 3C) can be drawn.

The amplitude of the cantilever oscillations is very large for small lift heights, and the images fade at separations larger than 80 nm. This is consistent with previous reports of a strong dependence of the tip-surface force on the vertical separation (13). More work needs to be done to understand this quantitatively. Note that some of the nanowires that appear in the topographic image are missing from the electric field image (Fig. 3). This is because either electrical contact to these nanowires has failed or electrical conduction along the wire length has been interrupted. The scanning force technique thus provides a

unique way of mapping the electrical properties of nanocomposites.

Applications of the metal nanowire composites include high-density electrical multi-feedthroughs and high-resolution plates for transferring a two-dimensional charge distribution between microelectronic devices. The semiconductor nanowires can be used in photodetector arrays of high spatial resolution, where each wire acts as a pixel of submicrometer dimensions. Also, with the application of the injection technique to ultrasmall channel insulators (channel diameter less than 50 nm) (16, 17), nanowire arrays can be made for fundamental studies of a variety of phenomena, such as quantum confinement of charge carriers and mesoscopic transport.

#### REFERENCES AND NOTES

1. An overview can be found in M. J. Yacaman, T. Tsakalakos, B. H. Kear, Eds., *Nanostruct. Mat.* 3 (nos. 1-6) (1993).
2. R. Roy, *Nanophase and Nanocomposite Materials*, S. Komarneni, J. C. Parker, G. J. Thomas, Eds. (Mat. Res. Soc. Symp. Proc. 286, Materials Research Society, Pittsburgh, PA, 1993), pp. 241-250, and references therein.
3. Early work on the preparation of ultrathin superconducting wires by injection of nanochannel matrices was done by W. G. Schmidt and R. J. Charles [*J. Appl. Phys.* 35, 2552 (1964)] and by V. N. Bogomolov [*Sov. Phys. Usp.* 21, 77 (1978)].
4. C. A. Huber and T. E. Huber, *J. Appl. Phys.* 64, 6588 (1988).
5. N. F. Borrelli and J. C. Luong, *Proc. SPIE* 866, 104 (1988).
6. B. L. Justus, R. J. Tonucci, A. D. Berry, *Appl. Phys. Lett.* 61, 3151 (1992).
7. M. J. Tierney and C. R. Martin, *J. Phys. Chem.* 93, 2878 (1989).
8. G. D. Stucky and J. E. Mac Dougall, *Science* 247, 669 (1990), and references therein.
9. P. M. Ajayan and S. Iijima, *Nature* 361, 333 (1993).
10. An array of parallel metal cylinders would be transparent to light that had a wavelength much larger than the cylinder diameter and separation and propagated along the cylinder axis [D. E. Aspnes, A. Heller, J. D. Porter, *J. Appl. Phys.* 60, 3028 (1986)].
11. Whatman Laboratory Division, Clifton, NJ.
12. J. W. Diggle, T. C. Downie, C. W. Goulding, *J. Electrochem. Soc.* 69, 365 (1969).
13. Y. Martin, D. W. Abraham, H. K. Wickramasinghe, *Appl. Phys. Lett.* 52, 1103 (1988).
14. J. E. Stern, B. D. Terris, H. J. Maimin, D. Rugar, *ibid.* 53, 2717 (1988).
15. Lift Mode operation, Digital Instruments, Inc., Santa Barbara, CA (patent pending).
16. R. C. Fumeaux, W. R. Rigby, A. P. Davidson, *Nature* 337, 147 (1989).
17. R. J. Tonucci, B. L. Justus, A. J. Campillo, C. E. Ford, *Science* 258, 783 (1992).
18. We thank V. Elings for valuable discussions, B. Schardt and S. Thedford for image processing, and S. Nourbakhsh for electron microscopy. This work was supported by the Army Research Office, the Independent Research Program of the Office of Naval Research, and the National Science Foundation.

22 October 1993; accepted 20 December 1993

## Green Fluorescent Protein as a Marker for Gene Expression

Martin Chalfie,\* Yuan Tu, Ghia Euskirchen, William W. Ward, Douglas C. Prasher†

A complementary DNA for the *Aequorea victoria* green fluorescent protein (GFP) produces a fluorescent product when expressed in prokaryotic (*Escherichia coli*) or eukaryotic (*Caenorhabditis elegans*) cells. Because exogenous substrates and cofactors are not required for this fluorescence, GFP expression can be used to monitor gene expression and protein localization in living organisms.

Light is produced by the bioluminescent jellyfish *Aequorea victoria* when calcium binds to the photoprotein aequorin (1). Although activation of aequorin in vitro or in heterologous cells produces blue light, the jellyfish produces green light. This light is the result of a second protein in *A. victoria* that derives its excitation energy

from aequorin (2), the green fluorescent protein (GFP).

Purified GFP, a protein of 238 amino acids (3), absorbs blue light (maximally at 395 nm with a minor peak at 470 nm) and emits green light (peak emission at 509 nm with a shoulder at 540 nm) (2, 4). This fluorescence is very stable, and virtually no photobleaching is observed (5). Although the intact protein is needed for fluorescence, the same absorption spectral properties found in the denatured protein are found in a hexapeptide that starts at amino acid 64 (6, 7). The GFP chromophore is derived from the primary amino acid sequence through the cyclization of serine-dehydrotyrosine-glycine within this hexapeptide (7). The mechanisms that produce the dehydrotyrosine and cyclize the poly-

M. Chalfie, Y. Tu, G. Euskirchen, Department of Biological Sciences, Columbia University, New York, NY 10027, USA.

W. W. Ward, Department of Biochemistry and Microbiology, Cook College, Rutgers University, New Brunswick, NJ 08903, USA.

D. C. Prasher, Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA.

\*To whom correspondence should be addressed.

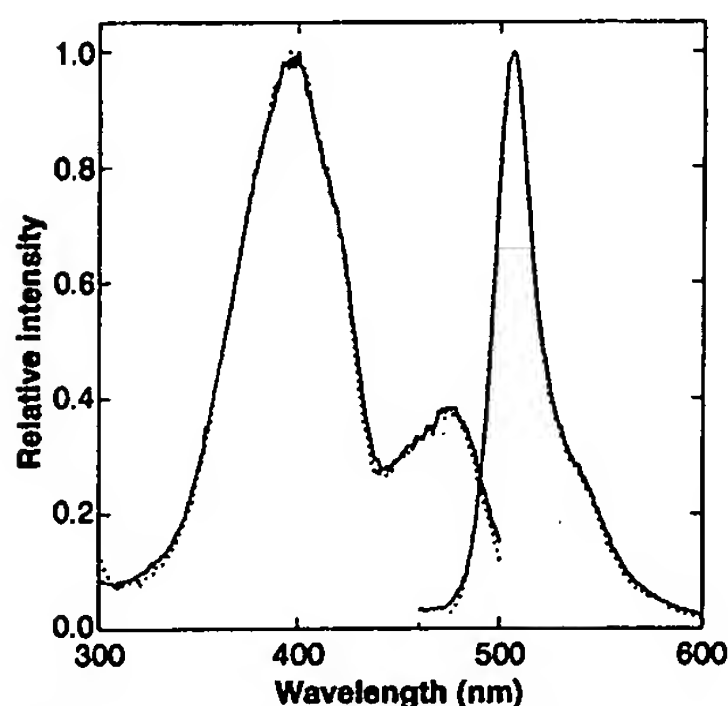
†Present address: U.S. Department of Agriculture, Building 1398, Otis Air National Guard Base, MA 02542, USA.



**Fig. 1.** Expression of GFP in *E. coli*. The bacteria on the right side of the figure have the GFP expression plasmid. Cells were photographed during irradiation with a hand-held long-wave UV source.

peptide to form the chromophore are unknown. To determine whether additional factors from *A. victoria* were needed for the production of the fluorescent protein, we tested GFP fluorescence in heterologous systems. Here, we show that GFP expressed in prokaryotic and eukaryotic cells is capable of producing a strong green fluorescence when excited by blue light. Because this fluorescence requires no additional gene products from *A. victoria*, chromophore formation is not species-specific and occurs either through the use of ubiquitous cellular components or by autocatalysis.

Expression of GFP in *Escherichia coli* (8) under the control of the T7 promoter results in a readily detected green fluorescence (9) that is not observed in control bacteria. Upon illumination with a long-wave ultraviolet (UV) source, fluorescent bacteria were detected on plates that contained the inducer isopropyl- $\beta$ -D-thiogalactoside (IPTG) (Fig. 1). Because the

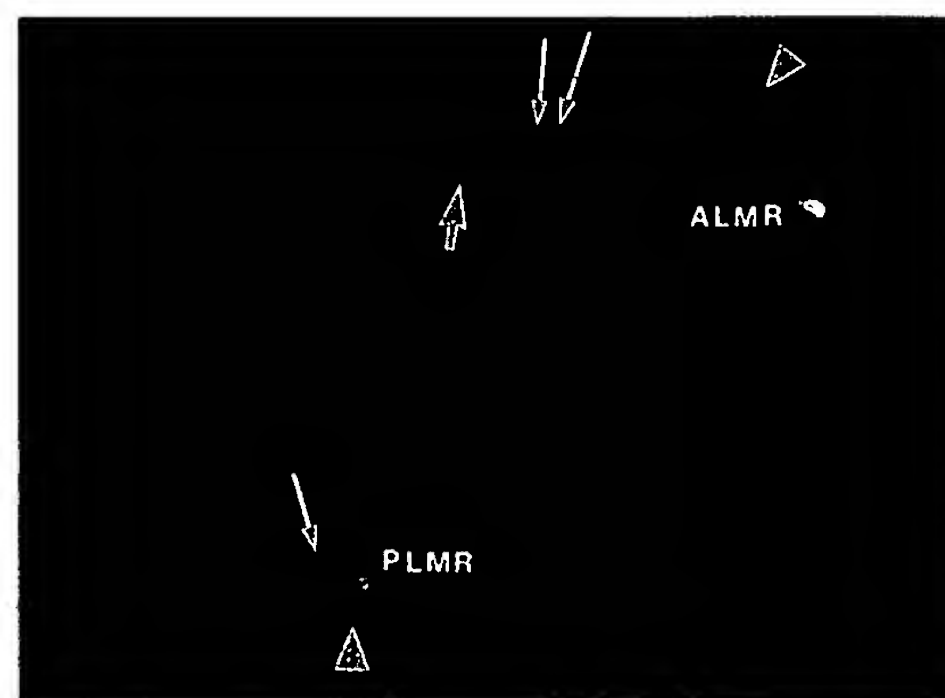


**Fig. 2.** Excitation and emission spectra of *E. coli*-generated GFP (solid lines) and purified *A. victoria* L form GFP (dotted lines).

cells grew well in the continual presence of the inducer, GFP did not appear to have a toxic effect on the cells. When GFP was partially purified from this strain (10), it was found to have fluorescence excitation and emission spectra indistinguishable from those of the purified native protein (Fig. 2). The spectral properties of the recombinant GFP suggest that the chromophore can form in the absence of other *A. victoria* products.

Transformation of the nematode *Caenorhabditis elegans* also resulted in the production of fluorescent GFP (11) (Fig. 3). GFP was expressed in a small number of neurons under the control of a promoter for the *mec-7* gene. The *mec-7* gene encodes a  $\beta$ -tubulin (12) that is abundant in six touch receptor neurons in *C. elegans* and less abundant in a few other neurons (13, 14). The pattern of expression of GFP was similar to that detected by MEC-7 antibody or from *mec-7-lacZ* fusions (13–15). The strongest fluorescence was seen in the cell bodies of the four embryonically derived touch receptor neurons (ALML, ALMR, PLML, and PLMR) in younger larvae. The processes from these cells, including their terminal branches, were often visible in larval animals. In some newly hatched animals, the PLM processes were short and ended in what appeared to be prominent growth cones. In older larvae, the cell bodies of the remaining touch cells (AVM and PVM) were also seen; the processes of these cells were more difficult to detect. These postembryonically derived cells arise during the first of the four larval stages (16), but their outgrowth occurs in the following larval stages (17), with the cells becoming functional during the fourth larval stage (18). The fluorescence of GFP in these cells is consistent with these previous results: no fluorescence was detected in these cells in newly hatched or late first-stage larvae, but fluorescence was seen in four of ten late second-stage lar-

**Fig. 3.** Expression of GFP in a first-stage *C. elegans* larva. Two touch receptor neurons (ALMR and PLMR) are labeled at their strongly fluorescing cell bodies. Processes can be seen projecting from both of these cell bodies. Halos produced from the out-of-focus homologs of these cells on the other side of the animal are indicated by arrowheads. The thick arrow points to the nerve ring branch from the ALMR cell (out of focus); thin arrows point to weakly fluorescing cell bodies. The background fluorescence is the result of the animal's autofluorescence.



vae, all nine early fourth-stage larvae, and seven of eight young adults (19). In addition, moderate to weak fluorescence was seen in a few other neurons (Fig. 3) (20).

Like the native protein, GFP expressed in both *E. coli* and *C. elegans* is quite stable (lasting at least 10 min) when illuminated with 450- to 490-nm light. Some photobleaching occurs, however, when the cells are illuminated with 340- to 390-nm or 395- to 440-nm light (21).

Several methods are available to monitor gene activity and protein distribution within cells. These include the formation of fusion proteins with coding sequences for  $\beta$ -galactosidase, firefly luciferase, and bacterial luciferase (22). Because such methods require exogenously added substrates or cofactors, they are of limited use with living tissue. Because the detection of intracellular GFP requires only irradiation by near UV or blue light, it is not limited by the availability of substrates. Thus, it should provide an excellent means for monitoring gene expression and protein localization in living cells (23, 24). Because it does not appear to interfere with cell growth and function, GFP should also be a convenient indicator of transformation and one that could allow cells to be separated with fluorescence-activated cell sorting. We also envision that GFP can be used as a vital marker so that cell growth (for example, the elaboration of neuronal processes) and movement can be followed in situ, especially in animals that are essentially transparent like *C. elegans* and zebra fish. The relatively small size of the protein may facilitate its diffusion throughout the cytoplasm of extensively branched cells like neurons and glia. Because the GFP fluorescence persists after treatment with formaldehyde (9), fixed preparations can also be examined. In addition, absorption of appropriate laser light by GFP-expressing cells (as has been done for Lucifer Yellow-containing cells) (25) could result in the selective killing of the cells.



## REFERENCES AND NOTES

- O. Shimomura, F. H. Johnson, Y. Saiga, *J. Cell. Comp. Physiol.* 59, 223 (1962).
- J. G. Morin and J. W. Hastings, *J. Cell. Physiol.* 77, 313 (1971); H. Morise, O. Shimomura, F. H. Johnson, J. Winant, *Biochemistry* 13, 2656 (1974).
- D. C. Prasher, V. K. Eckenrode, W. W. Ward, F. G. Prendergast, M. J. Cormier, *Gene* 111, 229 (1992).
- W. W. Ward, C. W. Cody, R. C. Hart, M. J. Cormier, *Photochem. Photobiol.* 31, 611 (1980).
- F. G. Prendergast, personal communication.
- O. Shimomura, *FEBS Lett.* 104, 220 (1979).
- C. W. Cody, D. C. Prasher, W. M. Westler, F. G. Prendergast, W. W. Ward, *Biochemistry* 32, 1212 (1993).
- Plasmid pGFP10.1 contains the Eco RI fragment encoding the GFP complementary DNA (cDNA) from *λgfp10* (3) in pBS(+) (Stratagene). The fragment was obtained by amplification with the polymerase chain reaction (PCR) [R. K. Saiki *et al.*, *Science* 239, 487 (1988)] with primers flanking the Eco RI sites and subsequent digestion with Eco RI. DNA was prepared by the Magic Mini-preps procedure (Promega) and sequenced (after an additional ethanol precipitation) on an Applied Biosystems DNA Sequencer 370A at the DNA sequencing facility at Columbia College of Physicians and Surgeons. The sequence of the cDNA in pGFP10.1 differs from the published sequence by a change in codon 80 within the coding sequence from CAG to CGG, a change that replaces a glutamine residue with arginine. [R. Heim, S. Emr, and R. Tsien (personal communication) first alerted us to a possible sequence change in this clone and independently noted the same change.] This replacement has no detectable effect on the spectral properties of the protein (Fig. 2). An *E. coli* expression construct was made with PCR to generate a fragment with an Nhe I site at the start of translation and an Eco RI site 5' to the termination signal of the GFP coding sequence from pGFP10.1. The 5' primer was ACAAGGCTAGCAAAGGAGAAGAAC and the 3' primer was the T3 primer (Stratagene). The Nhe I-Eco RI fragment was ligated into the similarly cut vector pET3a [A. H. Rosenberg *et al.*, *Gene* 56, 125 (1987)] by standard methods (26). The resulting coding sequence substitutes an Ala for the initial GFP Met, which becomes the second amino acid in the polypeptide. The *E. coli* strain BL21 (DE3) Lys S [F. W. Studier and B. A. Moffat, *J. Mol. Biol.* 189, 113 (1986)] was transformed with the resulting plasmid (TU#58) and grown at 37°C. Control bacteria were transformed with pET3a. Bacteria were grown on nutrient plates containing ampicillin (100 µg/ml) and 0.8 mM IPTG. [A similar PCR-generated fragment (11) was used in our *C. elegans* construct. As others are beginning to use pGFP10.1, we have heard that although similar PCR fragments produce a fluorescent product in other organisms (R. Heim, S. Emr, R. Tsien, personal communication; S. Wang and T. Hazelrigg, personal communication; L. Lanini and F. McKeon, personal communication) (23), the Eco RI fragment does not (R. Heim, S. Emr, R. Tsien, personal communication; A. Coxon, J. R. Chaillet, T. Bestor, personal communication). These results may indicate that elements at the 5' end of the sequence or at the start of translation inhibit expression.]
- We used a variety of microscopes (Zeiss Axio-phot, Nikon Microphot FXA, and Olympus BH2-RFC and BX50) that were equipped for epifluorescence microscopy. Usually, filter sets for fluorescein isothiocyanate fluorescence were used (for example, the Zeiss filter set used a BP450-490 excitation filter, 510-nm dichroic, and either a BP515-565 or an LP520 emission filter), although for some experiments filter sets that excited at lower wavelengths were used (for example, the Zeiss filter set with BP395-440 and LP470 filters and a 460-nm dichroic or with BP340-390 and LP400 filters with a 395-nm dichroic). In some instances, a xenon lamp appeared to give a more intense fluorescence than a mercury lamp when cells were illuminated with light around 470 nm, although usually the results were comparable. No other attempts were made to enhance the signal (for example, with low-intensity light cameras), although such enhancement may be useful in some instances. Previous experiments had shown that the native protein was fluorescent after glutaraldehyde fixation (W. W. Ward, unpublished data). S. Wang and T. Hazelrigg (personal communication) (23) have found that GFP fusion proteins in *Drosophila melanogaster* are fluorescent after formaldehyde fixation. We have confirmed that fluorescence persists after formaldehyde fixation with our *C. elegans* animals and with recombinant GFP isolated from *E. coli*. However, the chemicals in nail polish, which is often used to seal cover slips, did appear to interfere with the *C. elegans* GFP fluorescence.
- GFP was purified from 250-ml cultures of BL21 (DE3) Lys S bacteria containing TU#58; bacteria were grown in LB broth (26) containing ampicillin (100 µg/ml) and 0.8 mM IPTG. Induction was best when IPTG was present continually. Cells were washed in 4 ml of 10 mM Tris-HCl (pH 7.4), 100 mM NaCl, 1 mM MgCl<sub>2</sub>, and 10 mM dithiothreitol [A. Kumagai and W. G. Dunphy, *Cell* 64, 903 (1991)] and then sonicated (two times for 20 s each) in 4 ml of the same buffer containing 0.1 mM phenylmethylsulfonyl fluoride, pepstatin A (1 µg/ml), leupeptin (1 µg/ml), and aprotinin (2 µg/ml) and centrifuged at 5000 rpm for 5 min in the cold. The supernatant was centrifuged a second time (15,000 rpm for 15 min) and then diluted sevenfold with 10 mM Tris (pH 8.0), 10 mM EDTA, and 0.02% NaN<sub>3</sub>. Corrected excitation and emission spectra were obtained with a SPEX F1T11 spectrofluorometer (Metuchen, NJ) and compared with the purified L isoprotein form of GFP from *A. victoria* (M. Cutler, A. Roth, W. W. Ward, unpublished data). The excitation spectra were measured from 300 to 500 nm with a fixed emission wavelength of 509 nm, and the emission spectra were measured from 410 to 600 nm with a fixed excitation of 395 nm. All spectra were recorded as signal-reference data (where the reference is a direct measurement of the lamp intensity with a separate photomultiplier tube) at room temperature with 1-s integration times and 1-nm increments. The spectral band widths were adjusted to 0.94 nm for all spectra.
- Wild-type and mutant animals were grown and genetic strains were constructed according to S. Brenner [*Genetics* 77, 71 (1974)]. The plasmid pGFP10.1 was used as a template for PCR (with the 5' primer GAATAAAGCTAGCAAAGATGAG-TAAAG and the 3' T3 primer) to generate a fragment with a 5' Nhe I site (at the start of translation) and a 3' Eco RI site (3' of the termination codon). The DNA was cut to produce an Nhe I-Eco RI fragment that was ligated into plasmid pPD 16.51 (12, 27), a vector containing the promoter of the *C. elegans mec-7* gene. Wild-type *C. elegans* were transformed by coinjecting this DNA (TU#64) and the DNA for plasmid pRF4, which contains the dominant *rol-6* (*rol-6*) mutation, into adult gonads as described [C. M. Mello, J. M. Kramer, D. Stinchcomb, V. Ambros, *EMBO J.* 10, 3959 (1991)]. A relatively stable line was isolated (TU1710), and the DNA it carried was integrated as described by Mitani *et al.* (15) to produce the integrated elements *uls3* and *uls4* (in strains TU1754 and TU1755, respectively). Living animals were mounted on agar (or agarose) pads as described (16), often with 10 mM NaN<sub>3</sub> as an anesthetic (28) (another nematode anesthetic, phenoxypropanol, quenched the fluorescence) and examined with either a Zeiss universal or axiophot microscope. For *C. elegans*, a long-pass emission filter works best because the animal's intestinal autofluorescence (which increases as the animal matures) appears yellow (with band-pass filters the autofluorescence appears green and obscures the GFP fluorescence). Because much more intense fluorescence was seen in *uls4* than in *uls3* animals (for example, it was often difficult to see the processes of the ALM and PLM cells in *uls3* animals when the animals were illuminated with a mercury lamp), the former were used for the observations reported here. The general pattern of cell body fluorescence was the same in both strains and in the parental, nonintegrated strain (fluorescence in this strain was as strong as that in the *uls4* animals). The *uls4* animals, however, did show an unusual phenotype: both the ALM and PLM touch cells were often displaced anteriorly. The mature cells usually had processes in the correct positions, although occasional cells had abnormally projecting processes. These cells could be identified as touch receptor cells because the fluorescence was dependent on *mec-3*, a homeobox gene that specifies touch cell fate (13, 15, 18, 28). Expression of *mec-7* is reduced in the ALM touch cells of the head (but not as dramatically in the PLM touch cells of the tail) in *mec-3* gene mutants (13, 15). We find a similar change of GFP expression in a *mec-3* mutant background for both *uls3* and *uls4*. Thus, GFP accurately represents the expression pattern of the *mec-7* gene. It is likely that the reduced staining in *uls3* animals and the misplaced cells in *uls4* animals are results of either secondary mutations or the amount or position of the integrated DNA.
- C. Savage *et al.*, *Genes Dev.* 3, 870 (1989).
- M. Hamelin, I. M. Scott, J. C. Way, J. G. Culotti, *EMBO J.* 11, 2885 (1992).
- A. Duggan and M. Chalfie, unpublished data.
- S. Mitani, H. P. Du, D. H. Hall, M. Driscoll, M. Chalfie, *Development* 119, 773 (1993).
- J. E. Sulston and H. R. Horvitz, *Dev. Biol.* 56, 110 (1977).
- W. W. Walthall and M. Chalfie, *Science* 239, 643 (1988).
- M. Chalfie and J. Sulston, *Dev. Biol.* 82, 358 (1981).
- In adults, the thicker size of the animals and the more intense autofluorescence of the intestine tend to obscure these cells.
- These include several cells in the head (including the FLP cells) and tail of newly hatched animals and the BDU cells, a pair of neurons just posterior to the pharynx. Expression of *mec-7* in these cells has been seen previously (13, 15). The strongest staining of these non-touch receptor neurons are a pair of cells in the tail that have anteriorly directed processes that project along the dorsal muscle line. It is likely that these are the ALN cells, the sister cells to the PLM touch cells [J. G. White, E. Southgate, J. N. Thomson, S. Brenner, *Philos. Trans. R. Soc. London Ser. B* 314, 1 (1986)].
- The photobleaching with 395- to 440-nm light is further accelerated, to within a second, in the presence of 10 mM NaN<sub>3</sub>, which is used as a *C. elegans* anesthetic (11). However, when cells in *C. elegans* have been photobleached, some recovery is seen within 10 min. Further investigation is needed to determine whether this recovery represents de novo synthesis of GFP. Rapid photobleaching (complete within a minute) of the green product was also seen when *C. elegans* was illuminated with 340- to 390-nm light. Unlike the photobleaching with 395- to 440-nm light, which abolished fluorescence produced by the 340- to 390- or 450- to 490-nm light, photobleaching with 340- to 390-nm light did not appear to affect the fluorescence produced by 395- to 490- or 450- to 490-nm light. Indeed, the fluorescence produced by 450- to 490-nm light appeared to be more intense after brief photobleaching by 340- to 390-nm light. This selective photobleaching may indicate the production of more than one fluorescent product in the animal. These data on GFP fluorescence within *E. coli* and *C. elegans* are in contrast to preliminary studies that suggest that the isolated native and *E. coli* proteins are very photostable. We do not know whether this in vivo sensitivity to photobleaching is a normal feature of the jellyfish protein (the fluorescence in *A. victoria* has not been examined) or results from the absence of a necessary posttranslational modification unique to *A. victoria*.

- or from nonspecific damage within the cells.
22. Reviewed in T. J. Silhavy and J. R. Beckwith, *Microbiol. Rev.* 49, 398 (1985); S. J. Gould and S. Subramani, *Anal. Biochem.* 175, 5 (1988); and G. S. A. B. Stewart and P. Williams, *J. Gen. Microbiol.* 138, 1289 (1992).
  23. R. Heim, S. Emr, and R. Tsien (personal communication) have found that GFP expression in *Saccharomyces cerevisiae* can make the cells strongly fluorescent without causing toxicity. S. Wang and T. Hazelrigg (personal communication) have found that both COOH-terminal and NH<sub>2</sub>-terminal protein fusions with GFP are fluorescent in *D. melanogaster*. L. Lanini and F. McKeon (personal communication) have expressed a GFP protein fusion in mammalian (COS) cells.
  24. We have generated several other plasmid con-

- structions that may be useful to investigators. These include a pBluescript II KS (+) derivative (TU#65) containing a Kpn I-Eco RI fragment encoding GFP with an Age I site 5' to the translation start and a Bsm I site at the termination codon. Also available are *gfp* versions (TU#60 to TU#63) of the four *C. elegans lacZ* expression vectors (pPD16.43, pPD21.28, pPD22.04, and pPD22.11, respectively) as described (27) except that they lack the Kpn I fragment containing the SV40 nuclear localization signal.
25. J. P. Miller and A. Selverston, *Science* 206, 702 (1979).
26. J. Sambrook, E. F. Fritsch, T. Maniatis, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, ed. 2, 1989).

27. A. Fire, S. W. Harrison, D. Dixon, *Gene* 93, 189 (1990).
28. J. C. Way and M. Chalfie, *Cell* 54, 5 (1988).
29. We are indebted to A. Duggan and D. Xue for technical suggestions, to L. Kerr and P. Presley at the Marine Biological Laboratories at Woods Hole for help with microscopy, to M. Cutler and R. Ludescher for assistance in obtaining the excitation and emission spectra, to A. Fire for suggestions on vector construction, and to the colleagues listed in (8) and (23) for permission to cite their unpublished research. Supported by NIH grant GM31997 and a McKnight Development Award to M.C. and by American Cancer Society grant NP640 to D.C.P.

15 September 1993; accepted 16 November 1993

## RNA Polymerase II Initiation Factor Interactions and Transcription Start Site Selection

Yang Li, Peter M. Flanagan, Herbert Tschochner,\*  
Roger D. Kornberg†

An RNA polymerase II transcription system was resolved and reconstituted from extracts of *Schizosaccharomyces pombe*. Exchange with components of a *Saccharomyces cerevisiae* system was undertaken to reveal the factor or factors responsible for the difference in location of the transcription start site, about 30 base pairs and 40 to 120 base pairs downstream of the TATA box in *S. pombe* and *S. cerevisiae*, respectively. Two components, counterparts of human transcription factor IIF (TFIIF) and TFIH, could be exchanged individually between systems without effect on the start site. Three components, counterparts of human TFIIB, TFIIE, and RNA polymerase II, could not be exchanged individually but could be swapped in the pairs TFIIE-TFIH and TFIIB-RNA polymerase II, which demonstrates that there are functional interactions between these components. Moreover, exchange of the latter pair shifted the starting position, which shows that TFIIB and RNA polymerase II are solely responsible for determining the start site of transcription.

Synthesis of mRNA in eukaryotes requires RNA polymerase II and accessory factors, some which are general and act at most, if not all, promoters, and others of which confer specificity and control. Five general factors— $\alpha$ ,  $\beta$ ,  $\delta$ ,  $\epsilon$ , and  $\gamma$ —have been purified to homogeneity from the budding yeast *Saccharomyces cerevisiae* and have been identified as counterparts of human-rat factors TFIIE- $\epsilon$ , TFIH- $\delta$ , TFIID- $\tau$ , TFIIB- $\alpha$ , and TFIIF- $\beta\gamma$ , respectively (1–8). Because these factors assemble at a promoter in a complex with RNA polymerase II, interactions among them are assumed to be important for the initiation of transcription.

Most studies of general factor interactions have focused on binding (8). The results have shown that the order of assembly of the initiation complex on promoter DNA begins with factor d (TFIID), is followed by factor e (TFIIB), and then by polymerase and the remaining factors (6, 9). Factors  $\beta$  (TFIIF),  $\epsilon$ , and  $\gamma$  (TFIIF), however, bind directly to

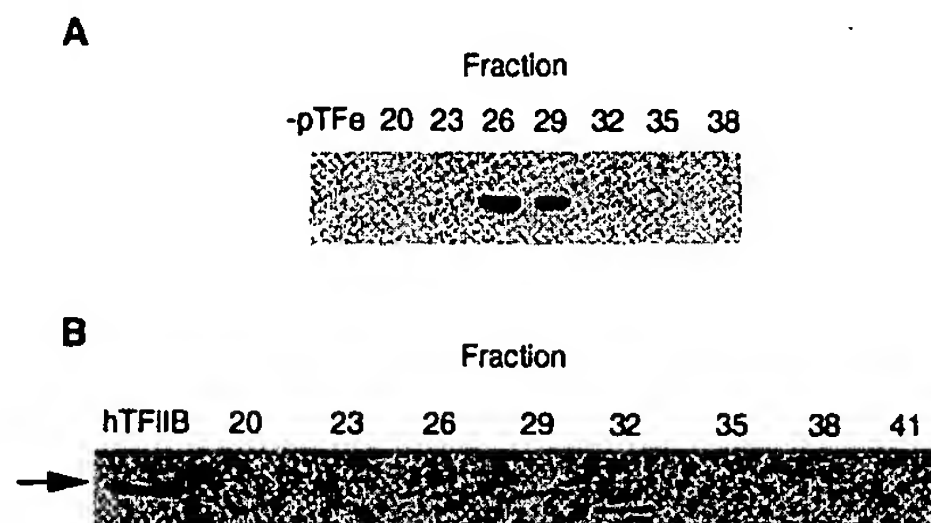
polymerase II, and as many as four of the five factors may assemble with the polymerase in a holoenzyme (10) before promoter binding. There are a couple of limitations implicit in these findings: The functional significance of interactions revealed by binding is questionable because only a few percent of initiation complexes give rise to transcripts, and there is little indication of the roles of the various interactions in the initiation process.

We have used a functional approach to

analyze general transcription factor interactions on the basis of the ability of factors to be exchanged between transcription systems. Exchange between *S. cerevisiae* and mammalian systems is of interest because of a marked difference in location of the transcription start site, 40 to 120 base pairs downstream of the TATA box in the former versus about 30 base pairs in the latter (11). The TATA-binding component (TBP) of factor d (TFIID) is functionally interchangeable between *S. cerevisiae* and humans (4, 12, 13), but the transcription start site remains characteristic of the particular transcription system, irrespective of the source of TBP. The factor or factors responsible for start site selection could not be identified by this approach because neither the other factors nor the polymerase proved interchangeable between *S. cerevisiae* and higher eukaryotic systems. We decided to use a *Schizosaccharomyces pombe* system because of its similarity to higher eukaryotes in the location of RNA polymerase II transcription start sites and its closer evolutionary relation to *S. cerevisiae*. Initiation from *S. pombe* promoters occurs about 30 base pairs downstream of the TATA box, and initiation from mammalian promoters introduced in *S. pombe* occurs at the same sites as in mammalian cells (14).

We have described the derivation of a chromatographic fraction from *S. pombe* that, upon addition of TBP, will support promoter-dependent RNA polymerase II transcription

**Fig. 1.** Factor e of *S. pombe* copurifies with a 35-kD polypeptide cross-reactive with human TFIIB antiserum. (A) Assay of fractions (2  $\mu$ l) from HAP (16) for pTFe activity. Assays were performed with the complete *S. pombe* system (16), except for the omission (first lane) of pTFe. (B) Immunoblot analysis of fractions (40  $\mu$ l) from HAP. Trichloroacetic acid precipitation, 12% SDS-polyacrylamide gel electrophoresis, and blotting onto nitrocellulose were followed by successive incubations with polyclonal human TFIIB antiserum (1:300) for 18 hours at 4°C and with goat monoclonal antibody to rabbit (1:2000) for 1 hour at 24°C as described (26). The 35-kD polypeptide (indicated by arrow) was the only cross-reactive species seen when smaller amounts of protein were loaded. Lane 1 contained 10 ng of human TFIIB.



Department of Cell Biology, Stanford University, School of Medicine, Stanford, CA 94305, USA.

\*Present address: Institute für Biochemie I, Der Universität Heidelberg, Im Neuenheimer Feld 328, 6900 Heidelberg, Germany.

†To whom correspondence should be addressed.